Kullback-Leibler Divergence(KLD) Formalism of The Stable Queue with KLD Applications to Biometrics

Dr Ismail A Mageed [*]

Article

Kullback-Leibler Divergence(KLD) Formalism of The Stable $M/G/1$ Queue with KLD Applications to Biometrics

Ismail A Mageed

drismail664@gmail.com

Abstract: The paper explores the Kullback-Leibler divergence formalism (KLDF) applied to the stable MG1 queue manifold. More potentially, both service time probability and cumulative functions which make KLDF exact are obtained. The credibility of KLDF is justified through consistency axioms. Some potential applications of Kullback-Leibler divergence to Biometry are presented. The paper concludes with closing remarks combined with some challenging open problems and the next phase of research.

Keywords: queue; server utilization (SU); short-range interactions; probability density function (PDF); Kullback-Leibler divergence formalism (KLDF); non-extensive maximum extropy (NME); extensive maximum entropy (EME); biometrics

I. Introduction

It has been common practice in probabilistic inverse approaches [1] to treat both measurable data and model parameters that are unknown as uncertain. This method provides deeper understanding of the uncertainty associated with the measured data and model parameters. [2–10]. KLD [11–16] is a method used to compare two probability distributions. In Probability and Statistics, when we need to simplify complex distributions or approximate observed data, KL Divergence helps us quantify the amount of information lost in the process of choosing an approximation. KLD measures the difference between the two distributions and assists us with comprehending the trade-off between accuracy and simplicity in statistical modelling.

Shannonian entropic measure [9,17], namely $H(p)$ reads as

$$H(p) = -\sum_{n=0}^{\infty} p(n)ln(p(n)) \tag{1}$$

to define "The minimum number of bits it would take us to encode our information".

KLD is commonly written as:

$$D_{KL}(p||r) = \sum_{n=0}^{\infty} p(n)ln\frac{p(n)}{r(n)}$$

We can precisely quantify lost when approximating one distribution with another using KL divergence.

This paper provides a roadmap of its contents, starting with some fundamental background in section I. The main results are given in section II. Section III deals with potential KLD applications to Biometrics. Finally, section IV provides conclusions, some emerging open problems, and future pathways of research.

According to [18–20], the maximum entropy state probability of the generalized geometric solution of a stable M/G/1 queue, subject to normalization, mean queue length (MQL), L and server utilization, ρ(<1) is given by:
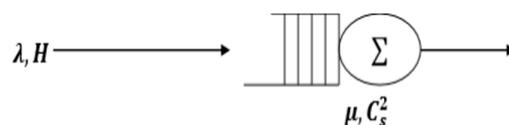
Figure 1. A Stable M/G/1 queue.

$$p(n) = \begin{cases} 1 - \rho, n = 0 \\ (1 - \rho)gx^n, n \geq 1 \end{cases} \quad (3)$$

where $g = \frac{\rho^2}{(L-\rho)(1-\rho)}, x = \frac{L-\rho}{L}$ and $L = \frac{\rho}{2}\left(1 + \frac{1+\rho C_s^2}{1-\rho}\right)$ (MQL for the underlying queue), ρ (= 1 - $p(0)$) and C_s^2 is the squared coefficient of variations). Obviously, $p(n)$ (c.f., (3)) reads:

$$p(n) = \begin{cases} 1 - \rho, n = 0 \\ \dfrac{2\rho(\frac{1+\rho\beta}{1-\rho} - 1)^{n-1}}{((\frac{1+\rho\beta}{1-\rho} + 1)^n)}, n \geq 1 \end{cases} \quad (4)$$

where $\beta = C_s^2$. The reader can observe the difference between our novel approach and that in (3) and (4). Moreover, it is notable that the newly obtained KL formalism in the current paper is more general which reduces to (3) as a special case.

[19] evaluated the credibility of Tsallis' NME formalism in terms of the four consistency axioms for large systems.

The credibility of KLDF NME formalism as a method of inductive reasoning is in Appendix A.

It has been shown (c.f., [19]) that the EME state steady probability of a stable M/G/1 queue that maximizes Shannon's entropy function (c.f.,[17]),

$$H(p_{1,s}) = -\sum_{n=0}^{\infty} p(n)\ln(p(n)) \quad (5)$$

with requirements,

Normalization,

$$\sum_{n=0}^{\infty} p(n) = 1 \quad (6)$$

SU,

$$p_{1,s}(0) = \sum_{n=0}^{\infty} h(n)p(n) = 1 - \rho, \rho = \frac{\lambda}{\mu} \quad (7)$$

where

$$h(n) = 1 \text{ for } n = 0 \text{ and } h(n) = 0 \text{ for } n = 1, 2, \ldots \quad (8)$$

P-K MQL,

$$<n> = \sum_{n=0}^{\infty} np(n) = \frac{\rho}{2}\left(1 + \frac{1+\rho C_s^2}{1-\rho}\right) \quad (9)$$

reads

$$p(n) = \begin{cases} p(0), n = 0 \\ p(0)\tau_s x^n \ n > 0 \end{cases} \quad (10)$$

where $p(0) = 1 - \rho, \tau_s = 2/(1 + C_s^2)$ and $x = \frac{<n> - \rho}{<n>}$.

II. MAIN RESULTS

Theorem 1. KLDF, namely, $p_{KL}(n)$, under (6)-(9) reads as

$$p_{KL}(n) = \begin{cases} p_{KL}(0) \ n = 0 \\ \dfrac{p_{KL}(0)\tau_s y^n}{q(0)} \ n > 0 \end{cases} \quad (11, KL)$$

where the initial state probabilities $p_{KL}(0)$ satisfy the SU constraint (6)

$$p_{KL}(0) = 1 - \rho \quad (12)$$

3

Such that

$$\tau_s = 2/(1 + C_s^2) \tag{13}$$

$$y = \frac{<n> - \rho}{<n>} = \frac{\rho}{\rho + (\tau_{KL}) p_{KL}(0)},$$

$$y^n = q(n) x^n, MQL = <n> = \frac{\rho}{2}\left(1 + \frac{\rho C_{s,KL}^2}{1-\rho}\right) \tag{14}$$

with

$$\frac{\rho(1-y)}{(1-\rho)y} = \frac{\tau_s}{q(0)} = \tau_{KL} \tag{15}$$

Proof

The Lagrangian, follows by maximizing KLDF under (6)-(9) to satisfy:

$$\left[\sum_{n=0}^{\infty} p_{KL}(n) \ln\left(\frac{p_{KL}(n)}{q(n)}\right) - \alpha\left(\sum_{n=0}^{\infty} h(n)\right) + (\beta - 1)\left(\sum_{n=0}^{\infty} 1\right) - \gamma\left(\sum_{n=0}^{\infty} n\right)\right] = 0 \tag{16}$$

Hence,

$$\ln\left(\frac{p_{KL}(n)}{q(n)}\right) + 1 + (\beta - 1) - \alpha h(n) + \gamma n = 0$$

Hence

$$p_{KL}(n) = q(n)(e^{-\beta})(e^{\alpha h(n)})(e^{-\gamma n}) \tag{17}$$

$n = 0, (17)$ translates to

$$p_{KL}(0) = q(0)(e^{-\beta})(e^{\alpha}) \tag{18}$$

Linking (17) with (18) implies

$$p_{KL}(n) = \frac{p_{KL}(0) q(n) \tau_s x^n}{q(0)},$$

$$x = e^{-\gamma} \in (0,1), \tau_s = e^{-\alpha} \in (0,1) \tag{19}$$

Hence, we have

Define, $y^n = q(n) x^n, y = \frac{<n> - \rho}{<n>}$, implies that $p_{KL}(n)$ will get the form:

$$p_{Kl}(n) = \begin{cases} p_{KL}(0) & n = 0 \\ \frac{p_{KL}(0) \tau_s y^n}{q(0)} & n > 0 \end{cases} \tag{20, KL}$$

Thus

$1 = p_{KL}(0) + p_{KL}(0)\tau_s \frac{y}{(1-y)q(0)}$, implying $\frac{\rho(1-y)}{(1-\rho)y} = \tau_{KL}$, which clearly implies $y = \frac{\rho}{\rho + (\tau_{Kl}) p_{KL}(0)}$. This proves the required results.

Clearly, from the above devised result

$$p_{KL}(n) = \begin{cases} p_{KL}(0) & n = 0 \\ \frac{p_{KL}(0) \tau_s y^n}{q(0)} & n > 0 \end{cases} \tag{20, KL}$$

It is known that $q(0) \in (0,1)$. it is evident that as $q(0) \rightarrow 1$, the KL formalism in (20,KL) will get the form:

$$p_{Sh}(n) = \begin{cases} p_{KL}(0) & n = 0 \\ p_{KL}(0) \tau_s y^n & n > 0 \end{cases}$$

which is the obtained Shannonian formalism in [20]. This shows the strength of the newly devised KLD formalism.

Theorem 2. The KL NME formalism, $p_{KL}(n)$, are exact when PDFs of service time read as

$$f_{s,KL}(t) = (1 - \tau_{KL}) u_0(t) + \mu(\tau_{KL})^2 e^{-\mu \tau_{KL}} \tag{21}$$

where $u_0(t)$ reads as

$$u_0(t) = \begin{cases} \infty, t = 0 \\ 0, t \neq 0 \end{cases} \tag{22}$$

such that

$$\int_{-\infty}^{\infty} u_0(t) = 1 \text{ and } \tau_s = 2/(1+ C_s^2), \ \tau_{KL} = \frac{\tau_s}{q(0)}$$

Proof

$Q_{KL}(z)(c.f., [20])$ for $p_{KL}(n) \ (c.f., 20, KL)$ reads as

$$Q_{KL}(z) = \sum_{n=0}^{\infty} p_{KL}(n)z^n, \ |z| < 1 \ (23)$$

Hence, by replacing $p_{KL}(n)$ of $(20, KL)$ into (23), we have

$$Q_{KL}(z) = \sum_{n=0}^{\infty} p_{KL}(n)z^n =$$

$$p_{KL}(0) + \frac{p_{KL}(0)\tau_{KL}yz}{(1-yz)} = \frac{p_{Kl}(0)(1-yz(1-\tau_{KL}))}{1-yz} \ (24)$$

Following [21],

$$Q_{Kl}(z) = \frac{p_{KL}(0)(1-z)\left(F_{s,KL}^*(\lambda-\lambda z)\right)}{F_{s,KL}^*(\lambda-\lambda z)-z} \ (25)$$

where

$$F_{s,Kl}^*(\theta) = E\left[e^{-\theta s}\right] = \int_0^{\infty} e^{-\theta t}f_{s,Kl}(t)dt \ (26)$$

Hence,

$$\frac{(1-z)\left(F_{s,KL}^*(\lambda-\lambda z)\right)}{F_{s,KL}^*(\lambda-\lambda z)-z} = \frac{(1-yz(1-\tau_{KL}))}{1-yz} \ (27)$$

Following (27), yields

$$F_{s,KL}^*(\lambda-\lambda z) = \frac{z((1-yz(1-\tau_{KL}))}{((1-yz(1-\tau_{KL}))-(1-z)(1-yz))} = \frac{(1-yz(1-\tau_{KL}))}{y\tau_{KL}-yz+1} \ (28)$$

Define, $\lambda - \lambda z = \theta$. Therefore,

$$z = 1 - \frac{\theta}{\lambda} \ (29)$$

Thus,

$$F_{s,KL}^*(\theta) = \frac{\mu\tau_{KL}+\theta(1-\tau_{KL})}{\theta+\mu\tau_{KL}} = (1-\tau_{KL}) + \frac{\mu(\tau_{KL})^2}{\theta+\mu\tau_{KL}} \ (30)$$

By inverting Laplace-StieltjesTransform, $F_{s,KL}^*(\theta)$ the GE$_{Kl}$-type *pdf* $f_{s,KL}(t)$(c.f., (21)) follows.

It is observed that as $q(0) \to 1$, $\tau_{KL} \to \tau_s$, which reduces to the Shannonian limiting case obtained (c,f., [20]).

Corollary 1. The CDF $F_{s,KL}(t)$o f the GE$_{KL}$ type of service time with the PDFs $f_{s,KL}(t)$ of Theorem 2 is captured fully by $F_{s,Kl}(t)$}, which reads as

$$F_{s,KL}(t) = 1 - \tau_{KL}e^{-\mu\tau_{KL}t} \ (31)$$

$\tau_s = 2/(1+ C_{s,}^2), \ \tau_{KL} = \frac{\tau_s}{q(0)}.$

Proof

We have

$F_{s,KL}(t) = \int_0^t f_{s,KL}(x) \, dx =$

$\int_0^t (1-\tau_{KL})u_0(x) \, dx + \mu\tau_{KL} \int_0^t e^{-\mu\tau_{KL}x} \, dx = (1-\tau_{KL}) + \frac{\mu\tau_{KL}}{\mu\tau_{KL}}(1-e^{-\mu\tau_{KL}t})$

$$= 1 - \tau_{KL}e^{-\mu\tau_{KL}t} \text{ QED (c.f., 31)}$$

For $q(0) \to 1$, the novel derivation (31) reduces to the formula in [20],

$F_s(t) = 1 - \tau_s e^{-\tau_s\mu t}$ with $\tau_s = \frac{2}{C_s^2+1}$.

Corollary 2. Following (21), we have

$$E(s_{KL}) = \frac{1}{\mu} \text{ (32)}$$

$$E(s_{KL}^2) = \frac{2}{\mu^2 \tau_{KL}} \text{ (33)}$$

$$C_{s,KL}^2 = \frac{E(s_{KL}^2)}{(E(S_{KL}))^2} - 1 = \frac{(2-\tau_{KL})}{\tau_{KL}} \text{ (34)}$$

$$\tau_{KL} = \frac{\tau_s}{q(0)}, \tau_s = 2/(1+ C_s^2).$$

Proof

The mean of S_{KL} is given by

$$E(S_{KL}) = \int_0^\infty t\, f_{s.KL}(t)dt = \int_0^\infty t\mu\tau_{KL}\, e^{-\mu\tau_{KL}t}dt = \mu\tau_{KL}\int_0^\infty t\, e^{-\mu\tau_{KL}t}dt \text{ (35)}$$

Introducing

$$\Gamma(m) = \int_0^\infty w^{m-1}\, e^{-w}dw$$

and substituting $w = \mu\tau_{KL}t$ on (35) and since $\Gamma(2) = 1$, it is implied that $E(s_{KL}) = \frac{\mu\tau_{KL}}{\mu^2(\tau_{KL})^2}\Gamma(2) = \frac{1}{\mu}$.

Moreover,

$$E(s_{KL}^2) = \int_0^\infty t^2\, f_{s,KL}(t)dt = \int_0^\infty t^2\mu\tau_{KL}\, e^{-\mu\tau_{KL}t}dt =$$

$$\mu\tau_{KL}\int_0^\infty t^2\, e^{-\mu\tau_{KL}t}dt \text{ (36)}$$

Let $w = \mu\tau_{KL}t,$, $E(s_{KL}^2)$ is given by (33). Hence $C_{s,q,R}^2$ by (34).

As $q(0) \to 1$, the new derivations (32)-(34) reduce to that in [20],

$$E(S) = \frac{1}{\mu} \text{ (37)}$$

$$E(S^2) = \frac{2}{\mu^2 \tau_{KL}} \text{ (38)}$$

$$C_{s,KL}^2 = \frac{E(S^2)}{(E(S))^2} - 1 = \frac{(2-\tau_{KL})}{\tau_{KL}} \text{ (39)}$$

with

$$\tau_{KL} = \frac{\tau_s}{q(0)}, \tau_s = 2/(1 + C_s^2).$$

III. KLD APPLICATIONS TO BIOMETRICS

The use of biometric information was addressed by [25], specifically accelerometer data collected by smartphones, for advanced authentication and on-line user identification. The proposed approach utilizes homological analysis to monitor the inherent walking patterns of different users in the accelerometer data. By transforming the expected persistence diagrams into probability distribution functions and measuring the discrepancy in walking patterns using the KLD, users can be identified with high accuracy.

In [25] a proposed user identification system that utilizes accelerometer data was employed. The raw data is then filtered, and the magnitude of the accelerometer readings is calculated. The system focuses on identifying the user's activity, specifically walking, by extracting the associated accelerometer data. This is illustrated by Figure 2.
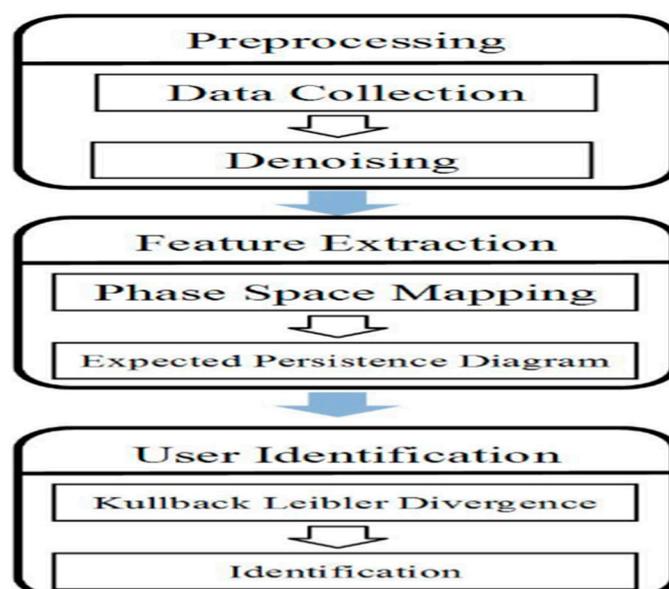
Figure 2. The diagram illustrates the flow of data processing for user identification, including dynamic segmentation and the extraction of reliable signal features for identification purposes[25].

The Sokoto Coventry Fingerprint Dataset (SOCOFing) [26] was used in this study to construct and evaluate different models. SOCOFing is a biometric fingerprint database consisting of 6,000 grayscale images collected from 600 African subjects. The dataset was partitioned into training, validation, and testing sets, and four neural network architectures were trained and evaluated, with the Res-WCAE model outperforming others in terms of noise reduction and achieving state-of-the-art performance. This can be seen from Figure 3 (c.f., [26]).
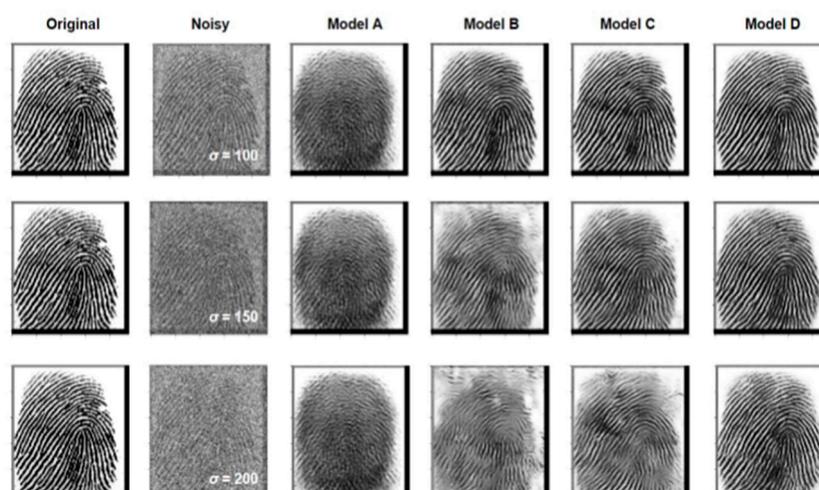
Figure 3. The samples shown depict original figures, noisy figures, and denoised figures for varying levels of noise.

To improve the ability of the proposed model [26] to make accurate predictions on unseen data, a regularized cost function that includes KLD regularization was introduced. This regularization is achieved by incorporating a prior distribution into the cost function.

Recently [27], LSR has been found to be effective in reducing the variation within a class by minimizing KLD between a uniform distribution and the predicted distribution of a neural network. This regularization method helps improve the performance of the network by providing more robust

and generalized predictions, particularly in classification tasks where reducing overconfidence and overfitting are important considerations.

To select the most informative sub-bands [28], the KL-divergence values of each dataset are normalized and averaged across the three datasets. Higher KL-divergence values indicate more discriminative sub-bands for classification. The sub-bands are chosen based on the highest average KLD values from all datasets, allowing for the identification of sub-bands that are discriminative across different datasets, as shown by Figure (c.f., [28]).
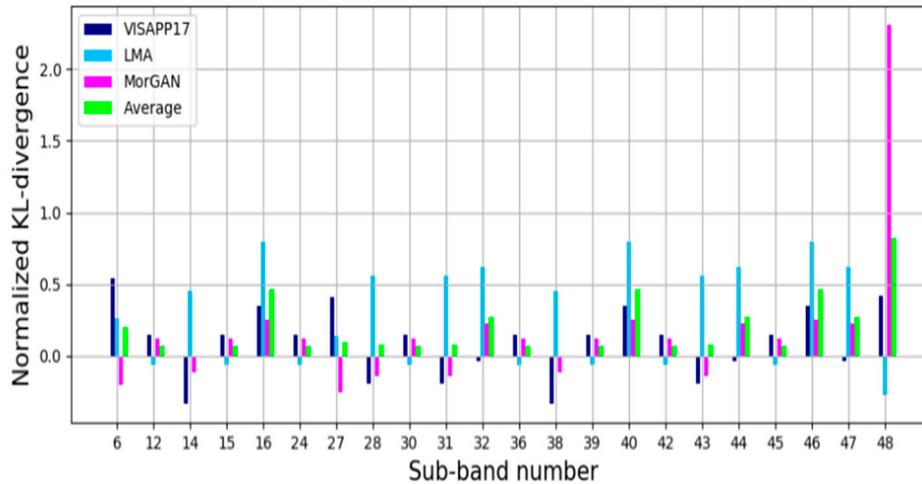
Figure 4. The average KL-divergence values in each sub-band for the three datasets, with the results displayed in green. KL-divergence is a measure of dissimilarity between probability distributions, and in this context, it is used to assess the discriminative power of different sub-bands in the datasets [28].

IV. SUMMARY, RESEARCH QUESTIONS COMBINED WITH NEXT PHASE RESEARCH

The ever-challenging problem of finding the closed form expression of the KLD formalism of the stable $M/G/1$ queue is solved in this paper. More fundamentally, the corresponding service time PDF and CDF for which the derived KLDF is exact are obtained. Some potential KLD applications to Biometrics are highlighted.

These are few emerging research questions:

- Can we unlock the mystery of the threshold of the obtained CDF(c.f., (31))?
- A really challenging open problem is replacing the proposed KLD in this paper by the corresponding KLD for Ismail's Entropy(IE) (c.f., [29]). This is a yet a great challenge to world mathematicians till current.
- Is the open problem of finding the info-geometric analysis of the derived KLDF solvable?
- Replacing the proposed version of KLD in this paper by the corresponding KLD of Ismail's entropy(IE)[29], can we get better results to advance Biometrics?

The next phase of research includes solving the above-listed open problems.

Appendix A. KL FORMALISM VS. EME CONSISTEN-CY AXIOMS

1. Uniqueness

It translates to, "If the same problem is solved twice in exactly the same way, the same answer is expected in both cases i.e., the solution should be unique" (c.f., [6]). Let f_{KL}, h_{Kl} be two PDFs such that:

$$H_{Kl}^*(f_{KL,N}) = H_K^*(h_{KL,N}) \text{ (A.1)}$$

Hence,

$$\sum_{n=1}^{N} \left(f_{KL,N} \right) ln(\frac{(f_{KL,N})}{q(N)}) = \sum_{n=1}^{N} \left(h_{KL,n} \right) ln(\frac{(h_{KL,N})}{q(N)}) \ (A.2)$$

By (A.2),it is implied that:

$$\left(f_{KL,N} \right)\left(ln\left(f_{KL,N} \right) - \ln\left(q(N) \right) \right) =$$
$$\left(h_{KL,N} \right)\left(ln\left(h_{KL,N} \right) - ln\left(q(N) \right) \right) \ (A.3)$$

Let the contradiction be true, namely, $\left(f_{KL,N} \right) \neq \left(h_{KL,N} \right)$. Thus $\exists\, \gamma > 1$ satisfying:

$$\left(f_{KL,N} \right) = \gamma\left(h_{KL,N} \right) \ (A.4)$$

Combining (A.3) and (A.4), we get

$$\gamma\left(h_{KL,N} \right)(ln\gamma + ln\left(h_{KL,N} \right) - ln(q(N)) =$$
$$\left(h_{KL,N} \right)(ln\left(h_{KL,N} \right) - ln(q(N)) \ (A.5)$$

Since $\left(h_{KL,N} \right)$ is non-zero, (A.5) can be transformed into

$$\gamma(ln\gamma) = (1 - \gamma)(ln\left(\frac{(h_{KL,N})}{q(N)}\right) \ (A.6)$$

By default, $q(N) \in (0,1)$.

By mathematical analysis, we have the following possibilities:

(1) $q(N) < h_{KL,N}$

It is implied that $\frac{(h_{KL,N})}{q(N)} > 1$, $ln(\frac{(h_{KL,N})}{q(N)}) > 0$.

This implies by (A.6), $\gamma > 1$, that $\gamma(ln\gamma) < 0$ (Contradiction)

(2) $q(N) = h_{KL,N}$

It is implied that $\frac{(h_{KL,N})}{q(N)} = 1$, $ln(\frac{(h_{KL,N})}{q(N)}) = 0$.

This implies by (A.6), $\gamma > 1$, that $\gamma(ln\gamma) = 0$ (Contradiction)

(3) $q(N) > h_{KL,N}$

It is implied that $\frac{(h_{KL,N})}{q(N)} < 1$, $ln(\frac{(h_{KL,N})}{q(N)}) < 0$.

Hence, KL divergence is negative, which contradicts that fact that KL divergence is non-negative(c.f.[22]).

Therefore., "there cannot be two distinct probability distributions $f_{kL,N}, h_{KL,N} \epsilon \Omega$ having the same KL divergence measure in Ω.Thus, KL formalism satisfies the axiom of uniqueness (c.f., [6]).

2. Invariance

The invariance axiom states that "The same solution should be obtained if the same inference problem is solved twice in two different coordinate systems" (c.f., [22]).Following the analytic methodology proposed in [6] and adopting the notation of Subsection 1, let Ξ be a coordinate transformation from state $\{S_n, n = 1,2, \dots, N\}$ to state $\{M_n, n = 1,2, \dots, N\}$, where M be a transformed set of N possible discrete states, namely $M = \{M_n, n = 1,2, \dots, N\}$ with $\Gamma(p_{Kl,N}(M_n) = \Xi^{-1}(p_{KL,N}(S_n))$, where J is the Jacobian $J = \frac{\partial(M_n)}{\partial(S_n)}$. Moreover, let $\Gamma\Omega$ be the closed convex set of all probability distributions Γ defined on M such that $\Xi\left(p_{Kl,N}(M_n)\right) > 0$ for all $M_n \epsilon M$, n = 1, 2, ...,N and $\sum_{n=1}^{N} \Xi\left(p_{KL,N}(M_n)\right) = 1$.It can be clearly seen that, transforming variables from $S_n \in S$ into $R_n \in R$, the extended KLD is transformation invariant [23] namely

$$H_{KL}^*\left(p_{KL,N} \right) = H_{KL}^*(\Xi\left(p_{KL,N}\right)) \ (A.7)$$

Thus, the EME formalism satisfies the axiom of invariance [6] since the minimum in $\Xi\Omega$ corresponds to the minimum in Ω" (c.f., [23]).

3. System Independence

It translates to "It should not matter whether one accounts for independent information about independent systems separately in terms of different probabilities or together in terms of the joint probability" (c.f., [6]). The joint probability for any independent systems Q and M is:

$$h_{Kl,N}(x_k, y_n) = Pr(X = x_k, Y = y_n) = f_{Kl,N}(x_k)\ g_{Kl,N}(y_n) \ (A.8)$$

Thus, KLD can be written as

$$H_{KL}^*\big(h_{Kl,N}\big) = \sum_{n=1}^{N}\big(h_{Kl,N}\big)\ln\left(\frac{\big(h_{Kl,N}\big)}{q(N)}\right) =$$

$$\sum_{n=1}^{N} f_{Kl,N}(x_k)\, g_{Kl,N}(y_n)\,\ln\left(\frac{\big(f_{Kl,N}(x_k)\, g_{Kl,N}(y_n)\big)}{q(N)}\right)\ (A.9)$$

Assume that

$$H_{KL}^*\big(h_{Kl,N}\big) = H_{KL}^*[f_{KL,N}] + H_{Kl}^*[g(Y_{KL,N})]\ (A.10)$$

By (A.9) and (A.10), this could be rewritten in the simpler form

$$fg\,ln\left(\frac{fg}{q(N)}\right) = f\,ln\left(\frac{f}{q(N)}\right) + g\,ln\left(\frac{g}{q(N)}\right)\ (A.11)$$

Hence,

$$fg\left(lnfg - lnq(N)\right) = f\,lnf + g\,lng - (f+g)lnq(N)\ (A.11)$$

Define $q(N) = 1,$ so (A.11) is

$$fg\,(lnfg) = f\,lnf + g\,lng\ (A.12)$$

(A.12) implies,

$$(fg)^{fg} = f^f g^g\ (A.13)$$

Let $f = g = \frac{1}{2}$ in (A.13)
Thus

$$(\tfrac{1}{4})^{\frac{1}{4}} = \tfrac{1}{2}\,(A.14)$$

(A.14) is impossible. Thus, system independence is defied because of long-range interactions.

4. Subset Independence(SI)

SI in a physical interpretation reads as "It does not matter whether one treats an independent subset of system states in terms of a separate conditional density or in terms of the full system density" (c.f., [24]).

In the given context, the notation and concepts related to an aggregate state of a system, denoted as x, and its associated probability distribution $f_{Kl}(x)$. The probability distribution represents the likelihood of the random variable X taking the value x. The text also mentions that the aggregate states ξ_i, where i ranges from 1 to L, can be expressed using this notation.

$$\sum_{S_i^*} f_{KL,i}\big(x_{ij}\big) = \xi_i\ (A.15)$$

We have

$$H_{KL}^*(f_{KL}) = \big(\sum_i \sum_{S_i} \xi_i\, f_{Kl,i}\big(x_{ij}\big)\ (A.16)$$

where $f_{Kl,i}(x)\,\epsilon\,\Omega.$ Equation (A.16) will read as

$$H_{Kl}^*(f_{KL}) = [\sum_i \xi_i \sum_{S_i} f_{KL,i}\big(x_{ij}\big)]\ )\ (A.17)$$

Thus,

$$H_{KL,i}^*\big(f_{Kl,i}\big) = \big(\sum_{S_i^*}\big(f_{Kl,i}\big)\big(x_{ij}\big)ln(\frac{\big(f_{Kl,i}\big)\big(x_{ij}\big)}{q(N)}))\ (A.18)$$

Hence,

$$H_{KL,i}^*\big(f_{Kl,i}(x_{ij})\big) = \left(\sum_{S_i^*} \ln\left(\left(\frac{\big(f_{Kl,i}\big)\big(x_{ij}\big)}{q(N)}\right)^{(f_{Kl,i})(x_{ij})}\right)\right)$$

$$= ln(\sum_{S_i^*}(\left(\frac{\big(f_{Kl,i}\big)\big(x_{ij}\big)}{q(N)}\right)^{(f_{Kl,i})(x_{ij})})\ (A.19)$$

Hence, apparently by the above proof it holds that:

$$\sum_{S_i^*} \left(\left(\frac{(f_{Kl,i})(x_{ij})}{q(N)} \right)^{(f_{Kl,i})(x_{ij})} > \left(\frac{(f_{Kl,i})(x_{ij})}{q(N)} \right)^{(f_{Kl,i})(x_{ij})} \right.$$

$$> \left(\frac{(f_{Kl,i})(x_{ij})}{q(N)} \right)^{q(N)} = \frac{(f_{Kl,i})(x_{ij})^{q(N)}}{(q(N))^{q(N)}} \ \text{(A.20)}$$

By (A.20), there exists a positive real number $0 < \sigma < 1$ satisfying:

$$\left(\frac{(f_{Kl,i})(x_{ij})}{q(N)} \right)^{(f_{Kl,i})(x_{ij})} = \frac{(f_{Kl,i})(x_{ij})^{q(N)}}{\sigma(q(N))^{q(N)}} \ \text{(A.21)}$$

Combining (A.19) and (A.21) implies

$$H_{KL,i}^*\big(f_{Kl,i}(x_{ij})\big) = \left(\sum_{S_i^*} \ln \left(\left(\frac{(f_{Kl,i})(x_{ij})}{q(N)} \right)^{(f_{Kl,i})(x_{ij})} \right) \right)$$

$$= \left(\frac{(f_{Kl,i})^{q(N)}(x_{ij})}{\sigma(q(N))^{q(N)}} \right) \text{(A.22)}$$

Following (A.23),

$$\big(f_{Kl,i}\big)^{q(N)}(x_{ij}) = \sigma\big(q(N)\big)^{q(N)} H_{KL,i}^*\big(f_{Kl,i}(x_{ij})\big) \ \text{(A.24)}$$

By (A.24),

$$f_{Kl,i}(x_{ij}) = q(N)\sigma^{\left(\frac{1}{q(N)}\right)} \big(H_{KL,i}^*\big(f_{Kl,i}(x_{ij})\big)\big)^{\left(\frac{1}{q(N)}\right)} \ \text{(A.25)}$$

Linking (A.16) with (A.25) yields

$$H_{KL}^*(f_{KL}) = \sum_i \xi_i q(N)\sigma^{\left(\frac{1}{q(N)}\right)} \big(H_{KL,i}^*\big(f_{Kl,i}(x_{ij})\big)\big)^{\left(\frac{1}{q(N)}\right)} \ \text{(A.26)}$$

(A.26) implies that KLD satisfies subset independence.

References

1. Z.S. Agranovich, and V.A. Marchenko, "The inverse problem of scattering theory," Courier Dover Publications; 2020.
2. H. Naman, N. Hussien, M. Al-dabag, and H. Alrikabi, "Encryption System for Hiding Information Based on Internet of Things," 2021, p. 172-183.
3. I.A.Mageed and Q.Zhang, "Formalism of the Rényian Maximum Entropy (RMF) of the Stable M/G/1 queue with Geometric Mean (GeoM) and Shifted Geometric Mean (SGeoM) Constraints with Potential GeoM Applications to Wireless Sensor Networks (WSNs)," electronic Journal of Computer Science and Information Technology, vol. 9, no. 1, 2023, p.31-40.
4. N. Mehr, M. Wang, M. Bhatt, M. Schwager, "Maximum-entropy multi-agent dynamic games: Forward and inverse solutions," IEEE Transactions on Robotics, 2023.
5. J. Korbel, "Calibration invariance of the MaxEnt distribution in the maximum entropy principle," Entropy, vol. 23, no. 1, 2021, p. 96
6. A. Golan and D.K. Foley, "Understanding the Constraints in Maximum Entropy Methods for Modeling and Inference," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 23, no. 3, 2022, p. 3994-3998.
7. D. Salvatore, "Theory and problems of statistics and econometrics," 2021.
8. A. Martini, S. Schmidt, and W. Del Pozzo, "Maximum entropy spectral analysis: A case study," arXiv preprint arXiv:2106.09499, 2021.
9. E. Pardo-Igúzquiza and P. A. Dowd, "Maximum entropy spectral analysis of uneven time series in the geosciences," Bol. Geol. Min, vol. 131, no. 2, 2020, p. 325-337.
10. I. A. Mageed and Q. Zhang, "An Introductory Survey of Entropy Applications to Information Theory, Queuing Theory, Engineering, Computer Science, and Statistical Mechanics, "In 2022 27th IEEE International Conference on Automation and Computing (ICAC), 2022, p. 1-6.
11. L. Sciullo, A.Trotta, and M. Di Felice, "Design and performance evaluation of a LoRa-based mobile emergency management system (LOCATE), "Ad Hoc Networks, vo. 96, 2020, p. 101993.

12. I.A. Mageed and Q. Zhang, "The Rényian-Tsallisian Formalisms of the Stable M/G/1 Queue with Heavy Tails Entropian Threshold Theorems for the Squared Coefficient of Variation, "electronic Journal of Computer Science and Information Technology, vol. 9, no. 1, 2023, p. 7-14.

13. I. A. Mageed, "The Entropian Threshold Theorems for the Steady State Probabilities of the Stable M/G/1 Queue with Heavy Tails with Applications of Probability Density Functions to 6G Networks, "electronic Journal of Computer Science and Information Technology, vol. 9, no. 1, 2023, p. 24-30.

14. I. A. Mageed, Q. Zhang, D. D. Kouvatsos, and N. Shah, "M/G/1 queue with Balking Shannonian Maximum Entropy Closed Form Expression with Some Potential Queueing Applications to Energy, "In IEEE 2022 Global Energy Conference (GEC), 2022, p. 105-110.

15. D.D. Kouvatsos et I. A. Mageed, "Formalismes de maximum d'entropie non extensive et inférence inductive d'une file d'attente M/G/1 stable à queues Lourdes, "Théorie des files d'attente 2: Théorie et pratique, 2021, p. 183.

16. I. A. Mageed and D. D. Kouvatsos, "The Impact of Information Geometry on the Analysis of the Stable M/G/1 Queue Manifold, "In ICORES, 2021, p. 153-160.

17. I. A. Mageed and Q. Zhang, "Inductive Inferences of Z-Entropy Formalism (ZEF) Stable M/G/1 Queue with Heavy Tails, "In IEEE 2022 27th International Conference on Automation and Computing (ICAC), 2022, p. 1-6.

18. I.A. Mageed and Q. Zhang, "Threshold Theorems for the Tsallisian and Rényian (TR) Cumulative Distribution Functions (CDFs) of the Heavy-Tailed Stable M/G/1 Queue with Tsallisian and Rényian Entropic Applications to Satellite Images (SIs)," electronic Journal of Computer Science and Information Technology, vol. 9, no. 1, 2023, p. 41-47.

19. I. A. Mageed, Q. Zhang, and B. Modu, "The Linearity Theorem of Rényian and Tsallisian Maximum Entropy Solutions of The Heavy-Tailed Stable M/G/1 Queueing System entailed with Potential Queueing-Theoretic Applications to Cloud Computing and IoT, "electronic Journal of Computer Science and Information Technology, vol. 9, no. 1, 2023, p. 15-23.

20. G. Malik, S. Upadhyaya, and R. Sharma, "Particle swarm optimization and maximum entropy results for $M^X/G/1$ retrial g-queue with delayed repair," International Journal of Mathematical, Engineering and Management Sciences, vol. 6, no. 2, 2021, p. 541.

21. B. Y. Lichtsteiner, "Delays in queues of queuing systems with stationary requests flows, "T-Comm-Телекоммуникации и Транспорт, vol. 15, no. 2, 2021, p. 54-58.

22. N. Anand and M. A. Saifulla, "A Probabilistic Method to Identify HTTP/1.1 Slow Rate DoS Attacks Check for updates," Communication and Intelligent Systems: Proceedings of ICCIS, vol. 2, 2023, Volume 2. 2023, p. 689-717.

23. D. D. Kouvatsos and and I . A. Mageed, "Non-Extensive Maximum Entropy Formalisms and Inductive Inferences of Stable M/G/1 Queue with Heavy Tails, "in "Advanced Trends in Queueing Theory, "vol. 2, March 2021, Vladimir Anisimov and Nikolaos Limnios (eds.), Books in 'Mathematics and Statistics', Sciences by ISTE & J. Wiley, London, UK.

24. S. E. Avgerinou, E. A. Anyfadi, G. Michas, and F. Vallianatos, "A Non-Extensive Statistical Physics View of the Temporal Properties of the Recent Aftershock Sequences of Strong Earthquakes in Greece, "Applied Sciences, vol. 13, no. 3, 2023, p. 1995.

25. K. Yan, L. Zhang, and H. C. Wu, "Advanced Homological Analysis for Biometric Identification Using Accelerometer, "IEEE Sensors Journal, vol. 21, no. 6, 2020, p. 7954-7963.

26. Y.Liang and W. Liang, "ResWCAE: Biometric Pattern Image Denoising Using Residual Wavelet-Conditioned Autoencoder, "arXiv preprint arXiv:2307.12255, 2023 .

27. Y. G. Jung, C. Y. Low, J. Park, and A.B. Teoh, "Periocular recognition in the wild with generalized label smoothing regularization, "IEEE Signal Processing Letters, vol. 27, 2020, p. 1455-1459.

28. P. Aghdaie, B. Chaudhary, S. Soleymani, J. Dawson, and N. M. Nasrabadi, "Detection of morphed face images using discriminative wavelet sub-bands, "In 2021 IEEE International Workshop on Biometrics and Forensics (IWBF), 2021, p. 1-6.

29. I. A. Mageed and Q. Zhang, "An Information Theoretic Unified Global Theory For a Stable M/G/1 Queue With Potential Maximum Entropy Applications to Energy Works," In IEEE 2022 Global Energy Conference (GEC), 2022, p. 300-305.